Perbandingan Metode Robust *Least Trimmed Square* Dengan Metode *Scale* Dalam Mengestimasi Parameter Regresi Linear Berganda Untuk Data Yang Mengandung Pencilan

Musafirah[1], Raupong[2], Nasrah Sirajang[3]

ABSTRAK

Metode estimasi yang banyak digunakan untuk mengestimasi parameter model regresi adalah metode kuadrat terkecil. Metode ini mempunyai asumsi-asumsi yang harus dipenuhi dalam penggunaan data riil. Akan tetapi, asumsi-asumsi tersebut terkadang dilanggar jika terdapat pengamatan yang bersifat pencilan. Hal ini akan mempengaruhi signifikansi. Oleh karena itu digunakan regresi *robust* untuk mengestimasi parameternya. Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi dari galat tidak normal atau adanya pencilan berpengaruh pada model. Regresi *Robust* memilki beberapa metode estimasi, dua diantaranya adalah metode estimasi *Least Trimmed Square* dan metode *Scale.* Kedua metode ini memilIki *breakdown point* yang tinggi terhadap adanya pengamatan yang bersifat pencilan dan memiliki algoritma yang lebih efisien dibandingkan dengan metode estimasi lainnya. Penerapan metode estimasi ini pada data nilai IPK terhadap nilai Ujian Nasional Mahasiswa jurusan matematika angkatan 2010 fakultas MIPA Universitas Hasanuddin, menghasilkan metode *Scale* memilki model yang lebih baik dibandingkan dengan metode estimasi *Least Trimmed Square*.

Kata kunci :*Ordinary Least Square, Regresi robust, Least Trimmed Square, Scale*

1.Pendahuluan

Regresi linear banyak digunakan dalam berbagai bidang hal analisis untuk melihat pengaruh suatu kondisi atau kejadian. Regresi linier merupakan metode statistika yang digunakan untuk membentuk model hubungan antara variabel terikat (*dependent;Y*) dengan satu atau lebih variabel bebas (*independent* ; *X*). Secara umum regresi linear terdiri dari dua yaitu regresi linear sederhana dan regresi linear berganda (Drapper dan Smith,1998).

Metode estimasi yang banyak digunakan untuk mengestimasi parameter model regresi adalah metode kuadrat terkecil *(Ordinary Least Square =OLS).* Metode ini mempunyai asumi-asumsi yang beberapa diantaranya dalam penggunaan data riil sering tidak dapat dipenuhi. Salah satu asumsi tersebut adalah mengenai kenormalan galat e_i yang sering dilanggar ketika adanya pengamatan yang bersifat pencilan. Akibat dari adanya pencilan, galat e_i tidak lagi berdistribusi normal atau variansi dari galatnya tidak lagi homogen. Dengan kondisi demikian, pengujian signifikansi parameter regresi selang kepercayaan akan menjadi tidak valid (Roesseeuw, dkk.1984).

Jika terdapat pencilan maka metode kuadrat terkecil tidak akurat untuk mengestimasi parameter. Untuk mengatasi masalah ini, salah satu metode yang digunakan adalah metode regresi *robust.* Metode ini dapat mengatasi pencilan dengan mencocokkan model regresi terhadap sebagian besar data. Suatu estimator *robust* mempunyai kemampuan mendeteksi pencilan sekaligus menyesuaikan estimasi parameter regresi.

Program Studi Statistika Jurusan Matematika
Fakultas MIPA

Metode *robust* estimasi *LTS* memiliki kemampuan yang lebih baik dibandingkan dengan metode-metode lainnya karena mampu mengatasi pencilan yang disebabkan oleh variabel bebas maupun variabel terikatnya. Selain itu metode *robust* estimasi *LTS* memilki algoritma yang lebih mudah dibandingkan metode lainnya karena dalam proses estimasinya, *LTS* hanya akan memangkas sebaran data berdasarkan jumlah pencilan yang teramati sehingga menghasilkan fungsi objektif yang mengecil dan konvergen ke 0 (Roesseuw, 1984). Sedangkan metode *robust* estimasi *S* juga merupakan metode dengan *high breakdown point* yang dikenalkan oleh Roesseuw dan Yohai (1984). Dengan nilai *breakdown* yang sama, metode ini memiliki efisiensi yang lebih tinggi dibanding estimasi *LTS* (Chen, 2002).

2. Tinjauan Pustaka

2.1 Regresi Linear Berganda

Regresi linear berganda adalah regresi linear yang terdiri dari satu variabel terikat dan lebih dari satu variabel bebas. Makna dari linear adalah linear dalam parameter dan variabelnya, yang berarti bahwa masing-masing parameternya hanya berpangkat 1 dan tidak dikalikan atau dibagi dengan parameter yang lain (Gujarati,1999).

Adapun bentuk persamaan regresi linear berganda adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i \quad ; \quad i = 1,2,\dots,n \qquad (1)$$

dimana :

Y_i : variabel terikat pada pengamatan ke-*i*

X_{ij} : variabel-variabel bebas pada pengamatan ke-*i* variabel ke- *j* (*j*=1,2,…,k)

β_0 : *intercept*

β_j : koefisien-koefisien regresi; *j*=1,2,…,k

ε_i : galat (*error*)

Model regresi berganda diatas dapat ditulis dalam bentuk matriks sebagai berikut :

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \qquad (2)$$

2.2 Metode Kuadrat Terkecil

Metode kuadrat terkecil (*Ordinary Least Square = OLS*) merupakan suatu metode untuk mendapatkan garis regresi yang baik yaitu sedekat mungkin dengan datanya sehingga menghasilkan prediksi yang baik (Widarjono, 2005).

Pada dasarnya, metode ini meminimumkan jumlah kuadrat error:

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \boldsymbol{\varepsilon}_i^2 = \boldsymbol{\varepsilon^t \varepsilon} = (\boldsymbol{y} - \boldsymbol{X\beta})^t(\boldsymbol{y} - \boldsymbol{X\beta})$$

$$\left.\frac{\partial \boldsymbol{S(\beta)}}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = -2\boldsymbol{X}^t\boldsymbol{y} + 2\boldsymbol{X}^t\boldsymbol{X}\widehat{\boldsymbol{\beta}} = 0$$

maka penaksir kuadrat terkecil dari $\boldsymbol{\beta}$ adalah

$$\widehat{\boldsymbol{\beta}} = [\boldsymbol{X}^t\boldsymbol{X}]^{-1}\,\boldsymbol{X}^t\boldsymbol{y} \qquad (3)$$

2.3 Pencilan
2.3.1 Definisi Pencilan

Dalam Soemartini (2007) pencilan didefinisikan dalam berbagai versi antara lain:

a. Menurut Ferguson (1961) pencilan adalah suatu data yang menyimpang dari sekumpulan data yang lain.
b. Menurut Bernet (1981) mendefinisikan pencilan sebagai pengamatan yang tidak mengikuti sebagian besar pola data dan terletak jauh dari pusat data.

2.3.2 Deteksi Pencilan

Ada beberapa cara untuk melihat atau menentukan apakah suatu pengamatan dapat dikategorikan sebagai pencilan, yaitu :

1. Untuk Pencilan pada variabel *X* menggunakan nilai *laverage* dengan persamaan:
$$\boldsymbol{H = X\,(X^tX)^{-1}\,X^t} \qquad (4)$$
Suatu pengamatan dikategorikan sebagai pencilan jika nilai $h_{ii} > \frac{2p}{n}$.

2. Untuk pencilan terhadap variabel *Y* adalah menggunakan *Studientized Deleted Residual (TRES)*. Menghitung statistik dengan Uji *TRES* yaitu :
$$TRES_i = \frac{e_i}{S_e} = e_i\left[\frac{n-k-1}{SSE(1-h_{ii})-e_i^2}\right]^{\frac{1}{2}} \qquad i = 1,2,.....,n \qquad (5)$$
dimana: $\varepsilon_i = Y_i - \hat{Y}_i$

S_e = simpangan baku galat

$\boldsymbol{h_{ii} = x_i\,(X^tX)^{-1}x_i^t}$

n = banyaknya pengamatan

Kriteria pengujian yang melandasi keputusan adalah:
$$|TRES_i| \begin{cases} > t_{\frac{\alpha}{2},n-k-1} & \text{, Tolak } H_0 \\ \le t_{\frac{\alpha}{2},n-k-1} & \text{, Terima } H_0 \end{cases}$$

3. Sedangkan untuk mendeteksi pencilan yang berkaitan dengan data berpengaruh maka digunakan ukuran $DFFITS_{(i)}$ sebagai berikut:
$$DFFITS_{(i)} = e_i\left[\frac{n-p}{SSE(1-h_{ii})-e_i^2}\right]^{\frac{1}{2}}\left[\frac{h_i}{(1-h_{ii})}\right]^{\frac{1}{2}}$$
Dimana :
ε_i = galat
SSE = jumlah kuadrat galat
$\boldsymbol{h_{ii}}$ = nilai *leverage.*

Suatu pengamatan dianggap berpengaruh (pencilan) jika nilai $DFFITS_{(i)} > 2\sqrt{\frac{p}{n}}$

2.4 Regresi Robust

Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi dari galat tidak normal atau adanya pencilan berpengaruh pada model (Ryan,1997). Metode ini merupakan alat penting untuk menganalisis data yang dipengaruhi oleh pencilan. Robust artinya parameter model tidak banyak berubah ketika sampel baru diambil dari populasi.

2.5 Estimasi *LTS (Least Trimmed Square)*

Estimasi *LTS* adalah dengan *high breakdown point* yang dikenalkan oleh Roesseuw (1984). *LTS* merupakan suatu metode estimator parameter regresi *robust* untuk meminimumkan jumlah kuadrat *h* residual (fungsi objektif).

$$\boldsymbol{E}^2{}_{LTS} = \sum_{i=1}^{h} e_{(i=n)}^2 \qquad (6)$$

dimana

h : $[n/2]+[(p+1)/2]$ (7)

$e_{(i)}^2$: Kuadrat gagal yang diurutkan dari yang terkecil ke yang terbesar

$$e_{(1)}^2 < e_{(2)}^2 < e_{(3)}^2 < \cdots e_{(i)}^2 < \cdots < e_{(h)}^2 < \cdots < e_{(n)}^2$$

Jumlah *h* menunjukkan sejumlah subset data dengan kuadrat fungsi objektif terkecil. Nilai h pada persamaan diatas akan membangun *breakdown point* yang besar sebanding dengan 50%.

Langkah-langkah yang dilakukan dalam mengestimasi parameter regresi *robust* metode *LTS:*

1. Mengestimasi koefisien regresi dengan MKT.
2. Menguji asumsi klasik analisis regresi linear.
3. Mendeteksi adanya pencilan dengan metode h_{ii}.
4. Tahap algoritma *LTS,* yaitu:
    a. Menghitung kuadrat *residual* e_i^2 dan menghitung h.
    b. Menghitung $E^2{}_{LTS}$.
    c. Melakukan estimasi parameter $b_{baru(i)}$ dari $h_{baru\,(i)}$ pengamatan.
    d. Menentukan kuadrat residual e_i^2 dari $h_{baru\,(i)}$ pengamatan.
    e. Menghitung $E^2{}_{LTS(baru)}$.
    f. Melakukan C-*steps* yaitu tahap d sampai f untuk mendapatkan fungsi objektif *(h)* yang terkecil dan konvergen ke 0.

2.6 Estimasi *S (Scale)*

Estimasi *S* pertama kali diperkenalkan oleh Rousseeuw dan Yohai (1984) merupakan estimasi *robust* yang dapat mencapai *breakdown point* hingga 50%. *Breakdown point* adalah ukuran umum proporsi dari pencilan yang dapat ditangani sebelum pengamatan tersebut mempengaruhi model. Karena estimasi *S* dapat mencapai *breakdown point* hingga 50% maka estimasi *S* dapat mengatasi setengah dari pencilan dan memberikan pengaruh yang baik bagi pengamatan lainnya.

Estimasi *S* didefinisikan

$$\hat{\boldsymbol{\beta}}^2{}_S = \min_\beta \hat{\sigma}_s\,(e_1, e_2, \ldots, e_n)$$

Langkah-langkah yang dilakukan dalam mengestimasi parameter pada regresi *robust* estimasi *S* adalah:

1. Mengestimasi koefisien regresi dengan MKT .
2. Menguji asumsi klasik analisis regresi linear.
3. Mendeteksi adanya pencilan pada data dengan metode *hii*
4. Langkah-langkah metode estimasi *S* ;
   a. Menghitung parameter $\hat{\beta}^0$ dengan MKT
   b. Menghitung nilai sisaan $e_i = y_i - \hat{y}_i$
   c. Menghitung nilai

$$\hat{\sigma}_s = \begin{cases} \dfrac{median\,|e_i - median\,(e_i)|}{0{,}6745} \\ \sqrt{\dfrac{1}{nK}\displaystyle\sum_{i=1}^{n} w_i e_i^2} \end{cases}$$

   d. Menghitung nilai $u_i = \frac{e_i}{\hat{\sigma}_s}$
   e. Menghitung pembobot w_i

$$w_i = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, & |u_i| \le c \\ 0, & |u_i| > c \end{cases}$$

   f. Menghitung parameter $\hat{\beta}_s$ dengan metode *(Weighted Least Square =WLS*) dengan pembobot w_i^0.
   g. Mengulangi langkah d sampai f hingga diperoleh nilai $\hat{\beta}_s$ yang konvergen.

3. Hasil dan Pembahasan

3.1 Estimasi Parameter Regrei Linear Berganda dengan Metode Kuadrat Terkecil

Konsep dari metode kuadrat terkecil biasa adalah mengestimasi parameter regresi dengan memininumkan jumlah kuadrat *error*. Untuk memperoleh estimator pada persamaan (1) maka dilakukan dengan metode kuadrat terkecil, yaitu :

$$\begin{aligned} \text{SSR} &= \textstyle\sum_{i=1}^{n} {e_i}^2 \\ &= {e_1}^2 + {e_2}^2 + \cdots + {e_n}^2 \\ &= [e_1 \quad e_2 \quad \ldots \quad e_n] \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \\ &= \boldsymbol{e}^t \boldsymbol{e} \end{aligned}$$

$$= (\boldsymbol{Y} - \boldsymbol{X\beta})^t(\boldsymbol{Y} - \boldsymbol{X\beta}), \qquad (8)$$

dengan menurunkan persamaan (8) terhadap $\boldsymbol{\beta}$ dan menyamakan hasil turunannya terhadap nol, maka diperoleh estimasi untuk $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X^tX})^{-1}\boldsymbol{X^tY} \qquad (9)$$

3.2 Estimasi Parameter Regresi

Data yang digunakan dalam penelitian ini merupakan data sekunder yang dikumpulkan melalui pengambilan sampel nilai Ujian Nasional dan Indeks Prestasi Kumulatif selama dua semester pada 50 orang mahasiswa jurusan Matematika Fakultas MIPA Universitas Hasanuddin angkatan 2010.

3.2.1 Pengujian Asumsi Klasik

Adapun Uji Asumsi Klasik yang dilakukan pada data nilai Ujian Nasional dan IPK Mahasiswa jurusan Matematika:

a. Uji Normalitas

Pada Uji Normalitas digunakan Rasio Skewness dan Kurtosis. Diperoleh bahwa rasio skewness = (-1.438)/0.481= (-2.99); sedang rasio kurtosis = 2.589/0.662= 3.91. Karena rasio skewness dan rasio kurtosis tidak berada diantara -2 hingga +2 maka dapat disimpulkan bahwa distribusi data tidak normal.

b. Uji Autokorelasi

Pada Uji Autokorelasi digunakan Durbin Watson dimana nilai durbin Watson pada data adalah 1,752 sehingga tidak terjadi autokorelasi

c. Uji heteroskedastisitas

Data tidak menunjukkan gejala heteroskedastisitas ditunjukkan oleh koefisien regresi dari masing-masing variabel bebas terhadap nilai absolute residualnya.

d. Uji Linearitas

Dengan menggunakan Analisis grafik pada data diperoleh model regresi yang terbentuk dinyatakan linear.

e. Uji Multikolinearitas

Dengan menggunakan cara partial Correlation bahwa terdapat beberapa nilai *Significance (2-tailed)* lebih kecul dari 0,05. Sehingga model regresi terbentuk dari data mengalami gejala multikolinearitas.

3.2.2 Estimasi Parameter Regresi menggunakan Metode Kuadrat Terkecil

Hasil estimasi parameter pada data diperoleh:

$$\hat{y_i} = 0{,}454 - 0{,}12\, X_1 - 0{,}05\, X_2 - 0{,}02\, X_3 + 0{,}146\, X_4 + 0{,}389\, X_5 - 0{,}04\, X_6 \qquad (10)$$

Dengan nilai R^2 = 19,4%. Yang artinya hanya 19,4% variabel bebas B.Indonesia (X_1), B.Inggris (X_2), Matematika (X_3), Fisika (X_4), Kimia (X_5), Biologi (X_6) yang mempengaruhi nilai IPK mahasiswa jurusan Matematika FMIPA UNHAS. Berdasarkan Anava terlihat bahwa niai Sig = 0,139 > α (0,05) yang berarti bahwa tidak ada hubungan antara variabel-variabel bebas dengan variabel terikatnya. Karena hanya variabel bebas Fisika (X_4) dan Kimia (X_5) yang korelasinya signifikan terhadap nilai IPK Mahasiswa jurusan Matematika maka estimasi parameter regresi hanya menggunakan variabel bebas Fisika (X_4) dan Kimia (X_5).

Dengan menggunakan dua variabel tersebut dilakukan pengujian asumsi klasik dan dperoleh tidak terdapat pelanggaran asumsi dari dua variabel tersebut, sehingga diperoleh estimasi parameter dengan menggunakan metode kuadrat terkecil:

$$\hat{y}_i = -1{,}183 + 0{,}131\ X_4 + 0{,}39\ X_5 \qquad (11)$$

Mendeteksi Pencilan

Dengan menggunakan data tersebut selanjutnya dilakukan pemeriksaaan pencilan terhadap variabel X nya.

Dari kriteria uji ini diperoleh beberapa pencilan yaitu data pada ke-5, ke-33

Hasil pemeriksaan pencilan

Data	Pengamatan ke-	h_{ii}	2p/n	Letak pencilan
1	5	0,13126	0,12	X
2	33	0.2704	0,12	X

3.2.3 Estimasi Parameter Menggunakan Metode *Least Trimmed Square.*

Dalam mengestimasi parameter regresi linear berganda menggunakan metode *Least Trimmed Square* pada data yang mengandung pencilan dilakukan langkah-langkah berikut:

a. Mengestimasi parameter dengan MKT

$$\hat{y}_i = -1{,}183 + 0{,}131\ X_4 + 0{,}39\ X_5$$

b. Menghitung kuadrat *residual* e_i^2 dan menghitung h.

Pada iterasi pertama digunakan model Regresi awal untuk menghitung kuadrat *residual* dan dari perhitungan tersebut dilakukan perhitungan terhadap h pengamatan,maka diperoleh h = 27, maka akan diurutkan nilai kuadrat *residual* dari yang terkecil ke yang terbesar dari data 1- 27

c. Menghitung E^2_{LTS}.

$$\boldsymbol{E}^2_{LTS} = 0{,}91394119$$

d. Melakukan estimasi parameter $b_{baru(i)}$ dari $h_{baru\ (i)}$ pengamatan.

$$\hat{y}_i = -\ 0{,}769 + 0{,}101\ X_4 + 0{,}384\ X_5$$

e. Menentukan kuadrat *residual* e_i^2 dari $h_{baru\ (i)}$ pengamatan.

Ditentukan *h baru* pengamatan, diperoleh h = 15 ,maka akan diurutkan kembali nilai *residual* dari yang terkecil ke yang terbesar sejumlah dengan *h baru* pengamatan

f. Menghitung E^2_{LTS}.

$$E^2_{LTS} = 1824{,}43149$$

g. Melakukan C-*steps* yaitu tahap d sampai f untuk mendapatkan fungsi objektif *(h)* yang terkecil dan onvergen ke 0.

Nilai Estimasi Parameter

Iterasi	h_i	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	E^2_{LTS}.
0	50	-1,183	0,131	0,39	
1	27	-0,769	0,101	0,384	0,91394119
2	15	-1,266	0,022	0,523	1824,43149
3	9	-0,936	0,074	0,395	7304126952
4	6	-5,677	-0,309	1,297	7,91417E+22
5	5	0,979	-0,253	0,48	8,01989E+48
6	4	-5,141	0,475	0,45	6,8833E+100
7	4	-5,141	0,475	0,45	5,73E+204

Langkah iterasi dengan Metode *robust* estimasi *Least Trimmed Square* di atas diperoleh nilai estimasi parameter pada data Nilai Ujian Nasional dan Nilai Indeks Prestasi Kumulatif Mahasiswa jurusan Matematika angkatan 2010 yang mengandung pencilan yaitu :

$$\hat{y}_i = -\ 5{,}141 + 0{,}475\ X_4 + 0{,}45\ X_5 \qquad (12)$$

3.2.4 Estimasi Parameter menggunakan Metode *Scale*

Dalam mengestimasi parameter selanjutnya ini, digunakan metode *Scale* dengan langkah- langkah sebagai berikut:

a. Menghitung parameter $\hat{\beta}^0$ dengan MKT

$$\hat{y}_i = -1{,}183 + 0{,}131\ X_4 + 0{,}39\ X_5$$

b. Menghitung nilai sisaan $e_i = y_i - \hat{y}_i$

c. Menghitung nilai Standar deviasi sisaan $\hat{\sigma}_s$

d. Menghitung nilai u_i menggunakan persamaan

e. Menghitung pembobot w_i

$$w_i \quad = \begin{cases} \left[1 - \left(\frac{\mathrm{u_i}}{\mathrm{c}}\right)^2\right]^2, & |\mathrm{u_i}| < \mathrm{c} \\ 0, & |\mathrm{u_i}| \geq \mathrm{c} \end{cases}$$

f. Menghitung parameter $\hat{\beta}_s$ dengan metode *WLS* dengan pembobot w_i^0.

Setelah melakukan tahap estimasi hingga diperoleh nilai w_i maka selanjutnya dilakukan estimasi parameter dengan menggunakan metode *WLS* sehingga diperoleh parameter $\hat{\beta}_s$

g. Mengulangi langkah b sampai f hingga diperoleh nilai $\hat{\beta}_s$ yang konvergen.

Nilai Standar deviasi dan parameter $\hat{\beta}_s$

Iterasi	$\hat{\sigma}_s$	Parameter $\hat{\beta}_s$		
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0	0,44255	-1,183	0,131	0,39
1	7,96405	1,162	0,09	0,162
2	10,3032	-0,993	0,122	0,379
3	10,3107	-1,116	0,127	0,387
4	10,3106	-1,120	0,127	0,387
5	10,3105	-1,121	0,127	0,387
6	10,3105	-1,121	0,127	0,387

Beberapa iterasi di atas diperoleh hasil dari metode *robust* estimasi *Scale* dalam mengestimasi data nilai Indeks Prestasi Kumulatif dan nilai Ujian Nasional Mahasiswa jurusan Matematika yang mengandung pencilan sebagai berikut:

$$\hat{y}_i = -1{,}121 + 0{,}127\, X_4 + 0{,}387\, X_5 \qquad (13)$$

3.2.5 Membandingkan Metode *Robust Least Trimmed Square* dan *Scale*

Iterasi metode *robust* estimasi *Least Trimmed Square* dan iterasi metode *robust* estimasi *Scale* diatas diperoleh nilai *Mean Square Error* untuk *Least Trimmed Square* = 0.640815, nilai *Mean Square Error* untuk *Scale* = 0.000386. Dengan melihat dari kedua nilai *Mean Square Error* dari kedua metode estimasi tersebut dapat disimpulkan bahwa dalam mengestimasi parameter pada data yang mengandung pencilan pada kasus Nilai Indeks Prestasi Kumulatif dan Nilai Ujian Nasional Mahasiswa jurusan Matematika Fakultas MIPA Universitas Hasanuddin metode *robust* estimasi *Scale* memiliki hasil yang lebih baik dibandingkan dengan *robust Least Trimmed Square* dan metode kuadrat terkecil.

4.Penutup

4.1 Kesimpulan

Hasil analisis yang telah dilakukan dan berdasarkan penjelasan yang telah diberikan , maka dapat diambil kesimpulan sebagai berikut.

1. Estimasi yang menggunakan metode kuadrat terkecil diperoleh model:
   $$\hat{y}_i = -1{,}183 + 0{,}131\, X_4 + 0{,}39\, X_5$$
   Estimasi yang menggunakan metode *robust Least Trimmed Square* diperoleh model :
   $$\hat{y}_i = -\, 5{,}141 + 0{,}475\, X_4 + 0{,}45\, X_5$$
   Dan estimasi yang menggunakan metode *robust* estimasi *Scale* diperoleh model:
   $$\hat{y}_i = -1{,}121 + 0{,}127\, X_4 + 0{,}387\, X_5$$
2. Estimasi parameter dengan metode *Least Trimmed Square* diperoleh nilai *mean square error* = 0,640815 dan pada metode *Robust Scale* diperoleh nilai *mean square*

error =0,000386. Hal ini menunjukkan bahwa metode *Robust Scale* lebih baik dalam mengestimasi parameter pada data yang mengandung pencilan.

4.2 Saran

Penelitian ini membahas tentang sifat-sifat dasar dan aplikasi dari estimator *Robust Least Trimmed Square* dan *Robust Scale* pada data Nilai Ujian Nasional dan Nilai Indeks Prestasi Kumulatif yang mengandung pencilan pada variabel bebasnya *(X)*. Untuk penelitian selanjutnya dapat dilakukan:

a. Penelitian estimasi parameter model regresi pada data yang mengandung pencilan pada variabel terikatnya *(Y)*.
b. Penelitian estimasi parameter model regresi pada data yang mengandung pencilan pada kedua variabelnya *(X* dan *Y)*.

DAFTAR PUSTAKA

Chen, Colin. 2002. *Robust Regression and Outier Detecttion with the Robustreg Procedure.* SUGI paper 265-267.SAS Insitute : Cary.NC.

Drapper,N.R dan Smith,H .1996. *Applied Regression Analysis,2nd edition.* New York : John Wiley & Sons. Chapman and Hall.

Gujarati, D. 1999. *Ekonometrika Dasar.* Terjemahan Sumarno Zain. Jakarta : Erlanngga.

Hogg, R.V, dan Allen T.C. 2005. *Introduction to mathematical Statistics*. USA: Pearson Prentice Hall.

Kristian,Yuddy.2010. Estimasi parameter model dalam regresi linear berganda dengan metode *LTS. Tesis*.Bandung: Universitas Padjadjaran.

Myers,Raymond H. 1989. *Classical and Modern Regression With Applications.* Boston: PWS-KENT.

Puput, Nuraidah. 2011. Estimasi Parameter Dalam Regresi Linear Berganda Dengan Metode *Least Median Square (LMS).* Makassar: Universitas Hasanuddin.

Roesseuw,P.J, dan A.M Leroy.1987. *Robust Regression and Outlier Detection.* New York: By John Wiley &Sons,Inc.

Roesseuw,P.J.1984. Least Median Squares Regression, *journal of the American Statistical Association.* Vol 79. Number 388.

Ryan,T.P. 1997. *Modern Regression Methods* . Canada: John Wiley & Sons,Inc.

Satman,Mehmet Hakan. 2013. *A* New Algoritm for Detecting Outlier in Linear Regression. *International Journal of Statistics and Probability;*Vol.2 No.3.. Turkey: University Istanbul.

Soemartini,2007. *Outlier* (pencilan).Bandung: UNPAD.

Zaman, A. Roesseuw, P.J, Orhan,M. 2001. Econometric Application of High Breakdown Robust Regression techniques. *Econometrics Letters.* Vol 71. 1-8.